Heredity (2006) 97, 291–295

www.nature.com/hdy

Patterns of microsatellite variation through a transition zone of a chromosomal cline in *Drosophila americana*

MA Schäfer[1], L Orsini[1], BF McAllister[1,2] and C Schlötterer[1]
[1]*Institut für Tierzucht und Genetik, Josef Baumann Gasse 1, A-1210 Wien, Austria;* [2]*Department of Biological Sciences, 143 Biology Building, University of Iowa, Iowa City, IA 52242-1324, USA*

Chromosomal rearrangements have been considered as important barriers to gene flow and were often used in the delineation of species. The original taxonomic designation of *Drosophila americana americana* and *Drosophila americana texana* is based on the presence/absence of a centric fusion between the X- and fourth chromosomes. *D. a. americana* presents the derived fused state, whereas *Drosophila a. texana* presents the freely segregating ancestral state. The degree of genetic separation between the two chromosomal forms is still controversial, with different genetic markers yielding contrasting results even when the same populations were analyzed. Using 27 polymorphic microsatellites, we re-evaluated patterns of genetic differentiation between six *D. americana* populations sampled through a transition zone of both chromosomal forms in the central United States. Our results clearly reject a scenario of two differentiated species forming a hybrid zone in a region of parapatry and indicate that gene flow minimizes genome-wide differentiation associated with the two chromosomal arrangements.
Heredity (2006) **97,** 291–295. doi:10.1038/sj.hdy.6800860; published online 5 July 2006

Keywords: genetic differentiation; gene flow; microsatellites; speciation; geographical separation; hybrid zone

Introduction

Chromosomal rearrangements, such as reciprocal translocations, inversions or chromosomal fusions, have been considered as important barriers to gene flow through their effect on recombination rather than on hybrid viability (Garagna *et al*, 1997; Noor *et al*, 2001; Rieseberg, 2001; Navarro and Barton, 2003). By suppressing recombination in heterozygous individuals, chromosomal rearrangements may act synergistically with isolation genes to diminish gene flow over larger genomic regions than would be otherwise possible and thus may promote population differentiation and even lead to speciation (Rieseberg, 2001; Navarro and Barton, 2003).

Drosophila americana americana and *Drosophila americana texana* have been the subject of much debate, depending on their classification as two different subspecies within the *Drosophila virilis* group (Throckmorton, 1982; Powell, 1997) or alternatively as two chromosomal forms of the same species (McAllister, 2002; Vieira *et al*, 2003). Both forms are native in the central to eastern United States. *D. a. texana* is the southern form, whereas *D. a. americana* is northerly distributed (Throckmorton, 1982). On the chromosomal level, *D. a. americana* is characterized by a derived fusion of the X- and fourth chromosomes, a fusion of Muller's elements A and B, respectively (Muller, 1940; Throckmorton, 1982), whereas *D. a. texana* retains the ancestral state, with the X- and the fourth chromosome segregating independently. Both chromosomal arrangements exhibit a broad geographical transition zone and the relative frequency of the arrangements is tightly associated with latitude. Fused chromosomes are present at high frequency in northern populations, whereas absent or at low frequency in southern populations (Vieira *et al*, 2001; McAllister, 2002). Apart from the chromosomal fusion, *D. a. americana* has been characterized by bearing a high frequency in chromosomal inversions that are absent or only present at low frequency in *D. a. texana* (Hsu, 1952). Therefore, recombination between alternative chromosomal forms is likely suppressed throughout a significant proportion of the euchromatic genome. For example, no recombinants were observed in the interval including the centromere and an inversion specific to the fused arrangement of the fourth chromosome (McAllister, 2003). The base of the X-chromosome also exhibits signs of reduced recombination (Vieira *et al*, 2003).

Despite extensive molecular work during the last decade, the degree of genetic separation between the two chromosomal forms remains controversial as different genetic markers yield contrasting results, even when the same populations were analyzed. Sequence data for several nuclear genes located on chromosomes X, 2, 3, and 4 indicated that flies with and without the X/4 fusion are indistinguishable at the DNA level (Hilton and Hey, 1996; Hilton and Hey, 1997; McAllister, 2002; Vieira *et al*, 2003), suggesting either high rates of ongoing gene flow or that the species have maintained a large effective population size, which resulted in a large number of shared ancestral alleles. Both hypotheses have been extensively discussed for *D. americana* (Hilton

Correspondence: C Schlötterer, Institut für Tierzucht und Genetik, Josef Baumann Gasse 1, A-1210 Wien, Austria.
E-mail: christian.schloetterer@vu-wien.ac.at
Received 13 March 2006; accepted 1 June 2006; published online 5 July 2006

and Hey, 1997; Vieira *et al*, 2003). More recently, absence of divergence between *D. a. americana* and *D. a. texana* has also been suggested by a phylogenetic analysis based on the mitochondrial sequences of Cytochrome b and Cytochrome c oxidase subunit II (Caletka and McAllister, 2004). Other studies, however, have detected significant population differentiation. Vieira *et al* (2001) found that amino-acid replacement polymorphisms at the *fused 1* gene, which is located near the base of the X-chromosome close to the fusion breakpoint, are significantly correlated with latitude and longitude in parallel with the chromosomal cline. These observations form the basis for the hypothesis that a balance between gene flow and divergent selection on the karyotypes themselves or on associated genes maintains a clinal distribution for the X/4 fusion (Vieira *et al*, 2001; McAllister, 2002). However, although this scenario might explain the significant differentiation between the two chromosomal forms at the *fused 1* locus as well as at some microsatellite loci mapping to chromosomal element B (Schlötterer, 2000), it cannot account for the genome-wide patterns of differentiation indicated by recent microsatellite analysis. Orsini *et al* (2004) detected significant genetic differences between populations of *D. a. americana* and *D. a. texana* at 11 out of 43 loci analyzed, and these loci were randomly distributed over the genome. These findings not only suggest that differentiation between *D. a. americana* and *D. a. texana* might be more pronounced than previously thought but also indicate that highly polymorphic microsatellites might be more powerful to resolve the genetic relationship between both cytological forms.

Most studies on the differentiation between *D. a. americana* and *D. a. texana* used population samples from the National *Drosophila* Species Resource Center (currently held at the Tucson Stock Center). These laboratory lines have been maintained in culture for many years and were represented by few strains sampled unevenly from a large geographical range in the United States. This sampling might have affected earlier findings in different manners. For instance, considering the small number of genes and strains analyzed, ancestral polymorphism might account for the lack of divergence between *D. a. americana* and *D. a. texana* detected in phylogenetic studies (Hilton and Hey, 1996, 1997). Significant differentiation between *D. a. americana* and *D. a. texana* detected by microsatellite analysis might be attributable to species differentiation, or alternatively to geographic separation as a consequence of uneven sampling (eg about half of the *D. a. americana* lines are derived from western locations in Montana and Nebraska). In this study, we re-evaluate patterns of microsatellite variation between *D. a. americana* and *D. a. texana* by analyzing six populations sampled through a transition zone of both chromosomal forms, in which the species delimitation is not clearly defined. These populations have been collected quite recently and have been used previously in studies of sequence differentiation at six nuclear genes (Vieira *et al*, 2001, 2003; McAllister, 2002). For the population genetic analysis presented here, we used 27 microsatellites randomly distributed over the genome. This is a significantly larger set of genetic markers than used in many past studies and represents a category of fast evolving DNA markers, which has been shown to be highly informative to infer phylogenetic relationships between closely related species.

Materials and methods

Fly strains

Flies originated from six localities collected along a broad longitudinal transect through the northern section of the range and a latitudinal transect through the transition zone where putative *D. a. americana* and *D. a. texana* populations adjoin in the central United States (Figure 1). The populations furthest apart from each other (OR01 and FP99) are separated by a distance of approximately 1000 km. The exact geographic sampling locations, the number of iso-female lines genetically analyzed as well as the relative frequency of fused–unfused chromosomes for each of the populations studied are given in Table 1. Collections were made in 1996, 1999, and 2001, with the digits in the population ID indicating the year of collection. Methods for examining chromosomal arrangement

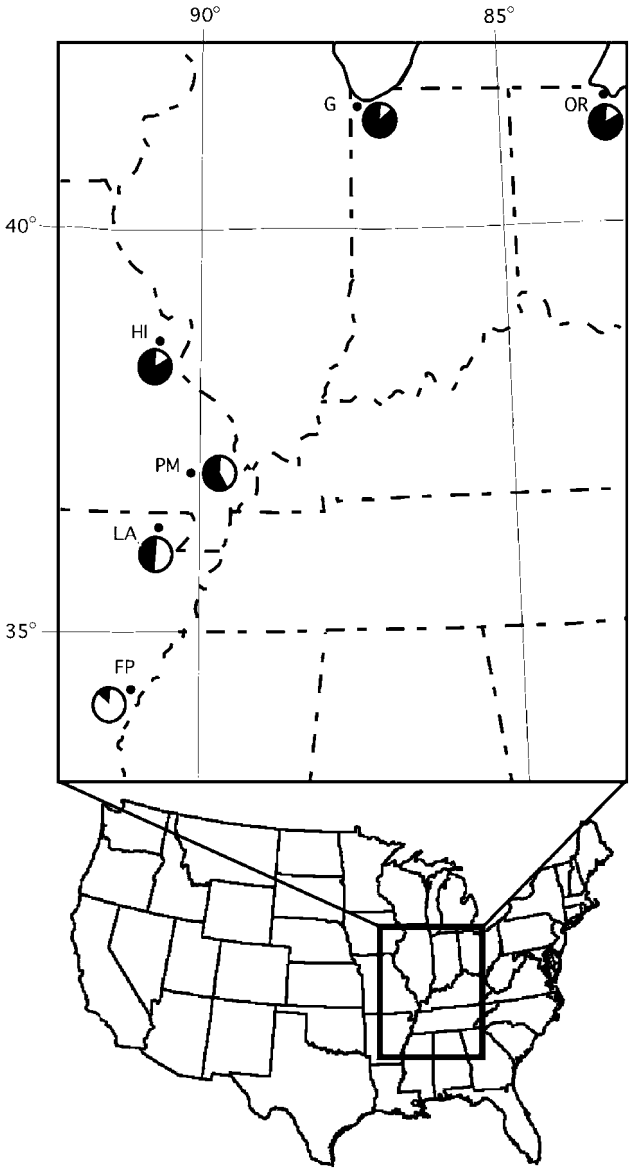

Figure 1 Geographic distribution and frequency of X/4 fusion in samples of *D. americana*. Each population contains both fused and unfused arrangements, and the black portion of the pie diagrams represents the frequency of the X/4 fusion.

Table 1 Populations analyzed with the corresponding number of individuals and the geographic location

Population	*No. of individuals*	*Location*	*Chromosomal state*
OR01	18	Ottawa NWR Toledo, Ohio	X/4 fusion 85.7%
G96	11	Gary, Indiana	X/4 fusion 98%
HI99	15	Howell Island, Missouri	X/4 fusion 84.6%
PM99	12	Puxico, Missouri	X/4 fusion 55.3%
LA99	10	Lake Ashbaugh, Arkansas	X/4 fusion 48.7%
FP99	19	Floodgate Park, Arkansas	X/4 fusion 13.6%

The percentage of X/4 fusion was determined previously (Vieira *et al*, 2001; McAllister, 2002) and is available at the website http://www.biology.uiowa.edu/mcallister/bfm_flies.html, edited by B McAllister.

and the frequencies of the alternative forms in these populations have been reported previously (McAllister and Charlesworth, 1999; McAllister, 2001, 2002). Both chromosomal forms are present in all populations. Each line was established from a single wild-caught female, and it has been subsequently maintained in laboratory culture at 18°C with transfers about once per month until individuals were processed in ethanol. For each iso-female line, one single female was genotyped.

Markers

We used a total of 27 polymorphic microsatellite markers to investigate patterns of genetic variation through the transition zone of both chromosomal forms. Most of the microsatellites were originally developed for *D. virilis* (Huttunen and Schlötterer, 2002) and *D. montana* (Orsini and Schlötterer, 2004) and successfully crossamplified in *D. americana*. Another set of six markers contained microsatellites previously isolated from P1 clones mapping to the Muller's chromosomal element B in *D. americana*, which is involved in the X/4 fusion (Schlötterer, 2000). DNA extraction and PCR amplification followed the protocols given in Schlötterer (2000); Huttunen and Schlötterer (2002); Orsini and Schlötterer (2004).

PCR-products were visualized by autoradiography, and allele sizes were determined running a 'PCR slippage ladder' together with a known size standard adjacent to the samples (Schlötterer and Zangerl, 1999). In order to evaluate the location of microsatellites on different chromosomes, we used a linkage map for *D. virilis* (Huttunen and Schlötterer, 2002) and a map for *D. montana* (MA Schäfer *et al*, unpublished data). Both species are close relatives of *D. americana* and thus should provide a reliable reference for the distribution of the microsatellites studied in the genome of *D. americana*.

Statistical analyses

We used two complementary statistical approaches to address the genetic separation between the alternative chromosomal arrangements. First, we calculated F_{ST} between populations (across loci and for each locus separately) using *F*-statistics according to Weir and Cockerham (1984). Statistical significance of F_{ST} values was calculated by permuting genotypes among populations for 10 000 times. This conservative procedure does not assume Hardy–Weinberg equilibrium and allows for linkage among the loci. The sequential Bonferroni correction procedure was applied to account for multiple testing (Sokal and Rohlf, 1995). All calculations were performed with version 3.12 of the MICROSATELLITE-ANALYZER software (Dieringer and Schlötterer, 2003). For original MSA input file, see Supplementary information.

Second, we addressed the potential differentiation between the two karyotypes within populations. As the karyotypes of individuals were not known and lines were highly inbred, we did this indirectly using a Bayesian clustering approach to detect cryptic population structure (Pritchard *et al*, 2000) as implemented in the computer program STRUCTURE version 2.1. The program determines the most probable number of homogeneous groups in a given data using multilocus genotypes and assigns individuals to one or more of them. We assumed prior values of the number of populations, *K*, from 1 to 6 and run 1 000 000 iterations of the MCMC, after a 'burn-in' period of 100 000 iterations. Three independent runs were performed to test the robustness of the results. Long runs were made to assure convergence of the MCMC chain and to obtain accurate estimates. For the simulations, we used a model without *a priori* consideration of the geographic origin of the samples that allows for admixture between populations.

For all analyses, we randomly selected one allele from each individual per locus to account for the inbreeding effect owing to the propagation of iso-female lines.

Results

We made two predictions concerning the genetic differentiation between the two karyotypes of *D. a. americana* and *D. a. texana*. If gene flow between individuals of different karyotype is suppressed, we predicted that (i) populations with different frequencies of fused or unfused chromosomes are differentiated from each other and that (ii) flies harboring fused and unfused chromosomes are genetically distinct within populations and are assignable to different groups.

Based on genotypic data of 27 microsatellites, however, we found that overall differentiation between populations was very low, albeit statistically significant (mean F_{ST} across loci $= 0.01$; 95% confidence interval from 0.003 to 0.017). Further analysis of pairwise F_{ST} values between populations (averaged over loci) showed that, whereas four of 20 comparisons were significant, only one remained statistically significant after sequential Bonferroni correction (Table 2). The high genetic similarity between *D. americana* populations appeared to be a genome-wide phenomenon, as none out of the 27 loci studied showed significant F_{ST} values after sequential Bonferroni correction (Table 3).

We performed a Bayesian analysis of population structure to test for 'cryptic subdivision', which might have remained undetected by the classical F_{ST} analysis. Previous analysis using representative lines of *D. a. americana* and *D. a. texana* showed that this Bayesian method has enough statistical power to distinguish between both putative 'species' (Orsini *et al*, 2004). Contrary to this expectation, however, we failed to recognize distinct genetic entities in our sample: a single cluster was supported with very high probability ($P > 0.999$).

Table 2 Pairwise F_{ST} values (Weir and Cockerham, 1984) among *D. americana* populations (lower triangular matrix) and corresponding *P*-values (upper triangular matrix; [a]significant after sequential Bonferroni correction)

Pairwise F_{ST}	*OR01*	*G96*	*HI99*	*PM99*	*LA99*	*FP99*
OR01	—	0.140	0.165	0.155	0.005	0.022
G96	0.011	—	0.144	0.799	0.095	0.529
HI99	0.007	0.010	—	0.626	0.000[a]	0.144
PM99	0.009	−0.010	−0.003	—	0.024	0.526
LA99	0.034	0.020	0.042	0.023	—	0.204
FP99	0.015	−0.001	0.007	−0.001	0.008	—

Table 3 *F*-statistics according to Weir and Cockerham (1984) for 27 microsatellite loci across six *D. americana* populations

Marker	*Chromosome*	*Global* F_{ST}	P-*value*
v11–23	2	0.018	0.096
vir6cs	2	0.027	0.144
msat21	2	−0.016	0.926
v10–10	2	0.026	0.100
v11–48.2	3	−0.027	0.908
msat19	3	0.021	0.047*
vir12cs	3	0.015	0.117
vir34	3	−0.017	0.870
vir84	3	−0.004	0.596
vir7	4	0.023	0.079
msat11	4	−0.006	0.615
mon35	4	0.028	0.095
mon37	4	0.026	0.072
v71.6.1	4*	0.023	0.110
v68.4	4*	0.039	0.040*
v.68.86.1	4*	0.016	0.150
v68.62	4*	−0.006	0.576
v68.86.2	4*	0.035	0.124
GPDH	4*	0.011	0.226
Msat4	5	−0.001	0.484
v71–03	X	−0.006	0.494
v68–06.1	X	−0.010	0.574
msat34	X	0.029	0.097
mon26	X	−0.018	0.884
mon31	X	0.011	0.367
mon6		0.026	0.089
vir47		0.000	0.452

Chromosomal location of markers marked with (*) was determined by chromosome extracted lines (Schlötterer, 2000) and that of the remaining ones from linkage maps of *D. virilis* (Huttunen *et al*, 2003) and *D. montana* (MA Schäfer *et al*, unpublished data).

Discussion

So far, molecular studies, which have addressed the genetic differentiation between *D. a. americana* and *D. a. texana*, provided ambiguous results leaving the genetic separation unresolved. We re-addressed the genetic differentiation between six *D. americana* populations sampled across a transition zone of the two chromosomal forms using 27 polymorphic microsatellite markers. Our results demonstrate that overall differentiation is low, albeit statistically significant. Nonsignificant differentiation is also supported by a model-based clustering method for multilocus genotype data indicating a single genetic neighborhood over the entire geographical area studied. These findings give insights into the role of recurrent gene flow *versus* shared ancestral polymorphism, the role of geographical separation *versus* species separation associated with the X/4th chromosomal fusion, and the evolutionary forces causing the chromosomal cline.

Inferring ongoing gene flow from molecular data can be significantly biased by ancestral variation, because both scenarios result in similar patterns of allele sharing. However, our study together with earlier microsatellite studies (Schlötterer, 2000; Orsini *et al*, 2004) provide strong evidence for ongoing gene flow across the transition zone of both chromosomal arrangements. The significant differentiation reported earlier (Schlötterer, 2000; Orsini *et al*, 2004) excludes that the low differentiation detected in the present study can be explained by shared ancestral polymorphism alone. Hence, gene flow must be significant. In line with this argument, heterozygous individuals regarding the X/4 fusion show normal fertility (Stone, 1949) and heterozygous individuals for the fusion have been collected in nature (Throckmorton, 1982; McAllister, 2002).

The considerable amount of gene flow through the transition zone further suggests that geographical separation rather than species separation associated with the X/4th chromosomal fusion is responsible for the differentiation between *D. a. americana* and *D. a. texana* detected earlier. Whereas in previous studies single representative lines from sparsely distributed geographical locations were investigated (Schlötterer, 2000; Orsini *et al*, 2004), our sampling was confined to the transition zone of the chromosomal cline including much smaller geographical distances among populations and these populations were represented by multiple lines. There was also at least some evidence for geographical separation influencing population differentiation based on the current study. Pairwise F_{ST} analysis showed four out of 20 comparisons between populations significant, out of which only one remained statistically significant after Bonferroni correction (Table 2). However, population differentiation was not associated with the cline of the X/4 fusion. These findings support the idea that geographical and not chromosomal separation may be important in shaping the population genetic structure in *D. americana*. It is highly unlikely that the discrepancy among studies results from the use of different microsatellite loci because a significant proportion of the loci were shared between the studies, including loci previously found significantly differentiated (Schlötterer, 2000; Orsini *et al*, 2004).

The two cytological forms of the X/4th chromosomal fusion in *D. americana* have been frequently regarded as two closely related species forming a hybrid zone in a restricted geographic area of parapatry. Theory on hybrid zones predicts the occurrence of steep allele frequency clines at many loci (Slatkin, 1973; Barton and Hewitt, 1989; Kruuk *et al*, 1999), which also applies to neutrally evolving DNA markers, such as microsatellites, owing to genetic hitchhiking (Maynard Smith and Haigh, 1974; Slatkin and Wiehe, 1998; Storz, 2005). Although the number of loci and the degree of differentiation may differ dependent on the strength of selection against heterozygous individuals, our data clearly reject the existence of a hybrid zone in a region of parapatry in *D. americana*. F_{ST} values estimated for individual loci were not significant, except for two, which is expected by chance alone (Table 3). Thus, the high genetic similarity of *D. americana* populations appears to be a genome-wide phenomenon and not confined to few specific blocks of DNA.

Furthermore, absence of significant differentiation at many neutral markers demonstrates that the chromosomal cline is unique relative to other segregating variations in the genome and is maintained despite the presence of gene flow. This pattern clearly refutes a scenario in which the chromosomal cline simply reflects neutral rates of dispersal through the region of secondary contact and strongly favors the hypothesis that spatially divergent selection on the karyotypes themselves or on associated genes, for example, the *fused 1* locus, maintains the geographical cline of the X/4 fusion in *D. americana* (Vieira *et al*, 2001; McAllister, 2002).

In conclusion, although chromosomal fusions/fissions have been reported as barriers to gene flow and have been often used in the delineation of species, the population genetic data presented here provide no evidence for genetic differentiation associated with the X/4th chromosomal fusion in *D. americana*. They rather support the idea that *D. americana* exists as a coherent species showing segregating polymorphism for a chromosomal fusion, which may be maintained by divergent selection along an ecological gradient.

Acknowledgements

We would like to thank Claus Vogl and members of the CS-lab for helpful discussions on earlier versions of the manuscript. This work was funded by the EU Research Training Network (HPRN-CT-2002–00266) and Fonds zur Förderung der wissenschaftlichen Forschung grants to CS and by the National Science Foundation under Grant No. DEB-0420399 to BFM.

References

Barton NH, Hewitt GM (1989). Analysis of hybrid zones. *Annu Rev Ecol Syst* **16**: 113–148.

Caletka BC, McAllister BF (2004). A genealogical view of chromosomal evolution and species delimitation in the *Drosophila virilis* species subgroup. *Mol Phylogenet Evol* **33**: 664–670.

Dieringer D, Schlötterer C (2003). Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes* **3**: 167–169.

Garagna S, Zuccotti M, Redi CA, Capanna E (1997). Trapping speciation. *Nature* **390**: 241–242.

Hilton H, Hey J (1996). DNA sequence variation at the period locus reveals the history of species and speciation events in the *Drosophila virilis* group. *Genetics* **144**: 1015–1025.

Hilton H, Hey J (1997). A multilocus view of speciation in the *Drosophila virilis* species group reveals complex histories and taxonomic conflicts. *Genet Res* **70**: 185–194.

Hsu TC (1952). *Chromosomal Variation and Evolution in the Virilis Group of Drosophila*. University of Texas Publication: Austin, TX.

Huttunen S, Aspi J, Hoikkala A, Schlötterer C (2003). QTL analysis of variation in male courtship song characters in *Drosophila virilis*. *Heredity* **92**: 263–269.

Huttunen S, Schlötterer C (2002). Isolation and characterization of microsatellites in *Drosophila virilis* and their cross species amplification in members of the *D. virilis*. *Mol Ecol Notes* **2**: 593–597.

Kruuk LE, Baird SJ, Gale KS, Barton NH (1999). A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. *Genetics* **153**: 1959–1971.

Maynard Smith J, Haigh J (1974). The hitch-hiking effect of a favorable gene. *Genet Res* **23**: 23–35.

McAllister BF (2001). Genetic analysis of sex-chromosome arrangement in *Drosophila americana*: a laboratory exercise for undergraduate of advanced placement students. *Dros Inf. Serv* **84**: 227–234.

McAllister BF (2002). Chromosomal and allelic variation in *Drosophila americana*: selective maintenance of a chromosomal cline. *Genome* **45**: 13–21.

McAllister BF (2003). Sequence differentiation associated with an inversion on the neo-X chromosome of *Drosophila americana*. *Genetics* **165**: 1317–1328.

McAllister BF, Charlesworth B (1999). Reduced sequence variability on the neo-Y chromosome of *Drosophila americana americana*. *Genetics* **153**: 221–233.

Muller HJ (1940). Bearings of '*Drosophila*'*work* on systematics. In: Huxley J (eds) *The New Systematics*. Clarendon Press: Oxford. pp 185–268.

Navarro A, Barton NH (2003). Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolut Int J Org Evolut* **57**: 447–459.

Noor MA, Grams KL, Bertucci LA, Reiland J (2001). Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci USA* **98**: 12084–12088.

Orsini L, Huttunen S, Schlötterer C (2004). A multilocus microsatellite phylogeny of the *Drosophila virilis* group. *Heredity* **4**: 1–5.

Orsini L, Schlötterer C (2004). Isolation and characterization of microsatellites in *Drosophila montana* and their cross-species amplification in *D. virilis*. *Mol Ecol Notes* **4**: 412–414.

Powell LR (1997). *Progress and Prospects in Evolutionary Biology: the Drosophila Model*. Oxford University Press: Oxford.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Rieseberg LH (2001). Chromosomal rearrangements and speciation. *Trends Ecol Evol* **16**: 351–358.

Schlötterer C (2000). Microsatellite analysis indicates genetic differentiation of the neo-sex chromosomes in *Drosophila americana americana*. *Heredity* **85**: 610–616.

Schlötterer C, Zangerl B (1999). The use of imperfect microsatellites for DNA fingerprinting and population genetics. In: Lubjuhn T (eds) *DNA Profiling and DNA Fingerprinting*. Basel: Birkhäuser. pp 153–165.

Slatkin M (1973). Gene flow and selection in a cline. *Genetics* **75**: 733–756.

Slatkin M, Wiehe T (1998). Genetic hitch-hiking in a subdivided population. *Genet Res* **71**: 155–160.

Sokal RR, Rohlf FJ (1995). *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd edn. WH Freeman and company: New York.

Stone WS (1949). The survival of chromosomal variation in evolution. In: Patterson JT (eds) *Studies in the Genetics of Drosophila*. University of Texas Publication: Austin, TX. pp 18–21.

Storz JF (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol Ecol* **14**: 671–688.

Throckmorton LH (1982). *The Genetic and Biology of Drosophila*. Academy Press: London.

Vieira CP, Coelho PA, Vieira J (2003). Inferences on the evolutionary history of the *Drosophila americana* polymorphic X/4 fusion from patterns of polymorphism at the X-linked paralytic and elav genes. *Genetics* **164**: 1459–1469.

Vieira J, McAllister BF, Charlesworth B (2001). Evidence for selection at the fused1 locus of *Drosophila americana*. *Genetics* **158**: 279–290.

Weir BS, Cockerham CC (1984). Estimating *F*-statistics for the analysis of population structure. *Evolut Int J Org Evolut* **38**: 1358–1370.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)